



[12]发明专利申请公开说明书

[21]申请号 98104273.2

[43]公开日 1998年10月14日

[11]公开号 CN 1195937A

[22]申请日 98.1.19

[30]优先权

[32]97.2.19 [33]US[31]802,598

[71]申请人 国际商业机器公司

地址 美国纽约州

[72]发明人 S·A·卡姆贝尔 D·C·克利奥塔
L·D·龙 L·门第托
M·E·米雅克 小L·H·奥维尔拜
B·拉雅拉曼 A·J·斯塔格

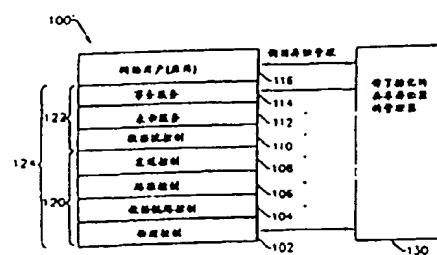
[74]专利代理机构 中国专利代理(香港)有限公司
代理人 邹光新 陈景峻

权利要求书 9 页 说明书 17 页 附图页数 12 页

[54]发明名称 通信系统中通信栈发送与接收数据的系统与方法

[57]摘要

信道中传输数据流前，通过提供数据分块的办法来作建立数据流的管理，从而减少数据的移动和拷贝，以提高通信系统性能。从通信栈更高一层接收的多个报头段被拷贝到数据流报头区内，则报头段顺序地存贮在数据流报头区内，具有指向报头区的入口记录的数据流缓冲器列表也产生了出来。指向数据段的入口记录也被存于数据流缓冲器列表。数据段在处理过程中就不用物理地移动或拷贝到数据流中，而由通信栈产生一个“虚拟”数据流来发送。



Best Available Copy

(BJ)第 1456 号

权 利 要 求 书

1. 一种方法，用来将一连串数据段和一连串报头段组合为数据流以供在网络信道中传输，每个上述数据段唯一地对应于一个上述的报头段，上述网络通道通常要求每个上述数据段相邻于与其唯一对应的上述报头段，使得在网络信道中传送数据段及相应的报头段更有效率，上述方法包括以下步骤：

5 将报头段数据中拷贝到数据流缓冲器中的数据流报头区，上述的报头段顺序地存入数据流缓冲器的上述数据流报头区；

10 将每个存在数据流缓冲器的报头区里的上述报头段同对应的一个上述数据段联系起来；

15 生成一个数据流缓冲器列表，其中包括在指向数据流缓冲器中的上述数据流报头区中的上述报头段的入口记录及指向上述数据段的入口记录。

其中上述数据段，上述数据流缓冲器列表，以及上述有数据流报头区的数据流缓冲器可以用来集中地形成在网络信道中传输的数据流，而不用在传送数据流之前物理地移动数据段或产生数据段的拷贝，上述数据流有报头区和数据区，这样，在上述报头区中的报头段可以同与其对应的上述的数据段不相邻地放入数据流中，上述数据流形成于这个的数据流在上述网络通道中作传送的时候。

20 2. 根据权利要求 1 的一种组合方法。其特征在于进一步包括下述步骤：

在网络信道中发送具有报头段的大量数据流报头区；

根据数据流缓冲器列表的指示，获取大量数据段；

25 根据数据流缓冲器列表的指示，将获取的大量数据段作为数据流数据区在网络通道中发送；

其中上述数据流报头区的报头段和上述数据区的数据段的结合形成了数据流，上述数据流报头区的报头段和与其对应的上述数据区的数据段不相邻地放在上述数据流内。

30 3. 根据权利要求 1 的一种组合的方法，其特征在于：上述的排列出现在具有多个层次的通信栈中，上述的拷贝步骤之前的步骤是，从通信栈的上述多个层中的一个层中接收数据段和报头段。

4. 根据权利要求 1 的一种组合的方法，其特征在于：上述的数据流缓冲器的报头区有一个块报头部分，上述的拷贝步骤之前的步骤是将数据流的特定信息存贮到上述的块报头部分。

5 5. 根据权利要求 1 的一种排列的方法，其特征在于相关步骤包括：

确定数据流中报头区的每个上述的报头段同与其相应的上述的数据段之间的偏移量；并且将所确定的偏移量存贮到与这个上述的报头段相关的数据流报头区内，这个上述的报头段对应于上述的数据段。

10 6. 根据权利要求 3 的一种组合方法，其特征在于上述通信栈就是在一种虚拟电信寻址方式（VTAM）系统中完成的通信栈。

15 7. 一种方法，在一个具有多层的通信栈中，用于从网络通道中接收不连续的数据流中得到的报头段及相应的数据段，上述接收的不连续数据流包含一个有许多报头段的报头区和一个有许多数据段的数据区，每个上述的数据段唯一地对应于一个上述的报头段，上述的其中一个数据段在特定的数据流中以一种不相邻的方式和一个与其对应的上述报头段相联系，这种方式包括以下步骤：

生成一个数据流缓冲器列表，其中包括一个指向上述不连续数据流中的一个上述报头段的报头入口记录，和一个指向不连续数据流中的一个上述数据段的数据入口记录，该数据入口记录唯一地与一个上述报头段对应；

20 存贮一个上述的报头段到一个报头段缓冲器内；

存贮一个上述的数据段到一个数据段缓冲器内；

25 将上述数据段中的一段及与其唯一对应的上述报头段中的一段在上述的信道中不相邻地传送到在上述的通信栈中的层，其中每个上述数据段中的一段和与其相关的上述报头段中的一段不相邻地传送到第二层而不需物理地移动与上述报头段中的一段相关的上述数据段中的一段，也不需对及与报头段中的一段相关的上述的数据段中的一段产生拷贝。

30 8. 根据权利要求 7 的一种组合的方法，其特征在于所产生步骤包括如下步骤：

识别不连续数据流中报头段中的一段；

将指向上述报头段的报头指针存贮到数据流缓冲器列表中；

识别出在不连续数据流中与已识别的上述报头段唯一对应的上述

一个数据段；

将指向上述数据段的已识别一段的数据指针存贮到数据流缓冲区列表中。

5 9.根据权利要求 7 的组合的一种方法，其特征在于，不连续传送步
骤包括以下步骤：

为一个上述数据段产生通信栈内部缓冲器列表，这个内部缓冲器列表唯一地与一个上述的数据段及与其关联的上述报头段相联；

在上述通信栈中将内部缓冲器列表从一个层次送到更高一个层
次，其中每个上述数据段和与其相关联的报头段不相邻地传送到通信栈
10 中更高的一层，而不需移动一个上述的数据段及与其相关的上述报头
段，也不需为与一个上述报头段相关联的一个上述的数据段产生一个拷
贝。

10 11.根据权利要求 9 的组合的一种方法，其特征在于产生一个通信栈
的内部缓冲器列表包括下述过程：

15 产生一个用于代表一个上述数据段的标记，这个标记唯一地与上述
数据段对应；

产生一个唯一地指向一个上述数据段的缓冲区列表报头指针，这个
上述的数据段与上述的报头段相对应；

将这个标记存贮在内部缓冲器列表中；

20 将缓冲器列表报头指针存贮到内部缓冲器列表中。

11.根据权利要求 7 的一种组合方法，其特征在于通信栈是一个完成
于虚拟电信寻址方式（VTAM）系统中的通信栈。

12.一种将大量数据段和大量报头段组合到一个数据流中以便于在
网系通道中传输的系统，每个上述的数据段唯一地对应于上述的报头
25 段，上述网络通道通常要求每个上述数据段和与其唯一对应的上述报头
段不相邻，这样就可以有效地在网络通道中传送数据段及相应的报头
段，该系统包括：

一个有数据流报头区的数据流缓冲器；

30 将大量报头段拷贝到数据流缓冲器的报头区中的装置，上述报头段
被顺序地存入数据流缓冲器中上述的数据流报头区中；

将每个存于数据流缓冲器的上述报头区的上述报头段和与其唯一
对应的上述数据段相联的装置；

用于产生数据流缓冲器列表的装置，包括指向数据流缓冲器中上述报头区的上述报头段的入口记录及指向上述数据段的入口记录；

其中上述的数据段，上述数据流缓冲器列表及有数据流报头区的上述数据流缓冲器都可以用来集中地形成用于在网络通道中传送的数据流，在传送数据流之前不需要物理地移动数据段，也不需要为数据段产生拷贝，上述的数据流具有报头区和数据区，这样，上述的报头区的报头段可以同与其对应的上述数据区中的数据段不相邻地放到数据流中，上述的数据流形成于在上述网络通道中传送该数据流的时候。

13.根据权利要求 12 的组合的系统，其特征在于更进一步包括：

10 在网络通道中传送含有大量报头段的数据流报头区的装置；

获取由数据流缓冲器列表指示的大量数据段的装置；

在网络通道中将获取的由数据流缓冲区列表所指示的大量数据段作为数据流数据区发送出去的装置；

其中上述数据流报头区中的报头段与上述数据流数据区的数据段的结合形成数据流，上述的数据流报头区的报头段和与其对应的上述数据流数据区的数据段不相邻地放在该数据流内。

14.根据权利要求 12 的一个用于组合的系统，与具有多个层次的通信栈有效地相联系的用于组合的系统，该系统还包括从通信栈的多个层次中的一层接收上述数据段和上述报头段的装置。

15.根据权利要求 12 的一个用于组合的系统，其特征在于上述数据流缓冲器的报头区有一个块报头部分，该系统还包括将数据流特有信息存贮到该块报头部分的方法。

16.根据权利要求 12 的一个用于组合的系统，其特征在于，用于相连的设备包括：

25 用于确定每个数据流中报头区的上述报头段和与其对应的数据段之间的偏移量的装置；

所确定用于将偏移量存贮到上述的报头区中的装置，该报头区和与一个上述数据段对应的上述报头段相关联。

17.根据权利要求 14 的一个用于组合的系统，其特征在于，通信栈是在虚拟电信寻址方式（VTAM）系统中实现的通信栈。

30 18.在通信栈中用于组合的一个系统具有多层报头段及相应的数据段，这些报头段和数据段在从网络通道中接收到的非相邻的数据流中，

该接收到的非相邻的数据流包括一个具有大量报头段的报头区和一个具有大量数据段的数据区，每个数据段唯一地与一个报头段对应，一个数据段和与其对应的特定的报头段以不相邻的方式存贮于该数据流中，该系统包括：

5 产生数据流缓冲器列表的装置，该缓冲器列表有一个指向非相邻数据流中一个上述报头段的报头入口记录，和一个指向唯一与该报头段对应的非相邻数据流中一个上述数据段的入口记录；

一个报头段缓冲器；

将一个上述的报头段存贮到一个报头段缓冲区内的装置；

10 一个数据段缓冲器；

将一个上述的数据段存储到一个数据段缓冲区内的装置；

将一个上述的数据段和一个与其第一对应的报头段不相邻地传递到该通信栈的第二层的装置，在该通信栈中一个上述的数据段同与其对应的上述的报头段不相邻地传送到第二层，而不需要物理地移动与上述报头段对应的数据段，也不需为及与上述报头段对应的数据段产生拷贝。

19.根据权利要求 18 的用于组合的系统，其特征在于产生装置包括：

用于在上述的不连续数据流中识别一个上述报头段的装置；

20 用于将所识别的上述报头段的报头指针存贮到数据流缓冲器列表中的装置；

用于在上述不连续的数据流中识别出与已识别的报头段唯一地对应的一个上述的数据段的装置；

25 用于将已识别的上述数据流的数据指针存贮到数据流缓冲器列表中的装置。

20.根据权利要求 18 的用于组合的一个系统，其特征在于，非相邻传送装置包括：

用于为一个上述数据段生成一个通信栈内部缓冲器列表的装置，这个内部缓冲器列表唯一地与一个上述的数据段及与其对应的报头段相关联；

30 用于将内部缓冲器列表从通信栈中一个层次送到更高层的装置。其中将上述的数据段和与其对应的报头段不相邻地送入该通信栈的更高

一层，而不需物理地移动与上述报头段相对应的上述的数据段，也不需为与上述报头段对应的上述的数据段产生拷贝。

21.根据权利要求 20 的用于组合的系统，其特征在于，产生内部缓冲器列表的装置包括：

5 用于产生代表一个上述数据段的标志的装置，该标志唯一地与相应
的数据段相联系；

用于产生唯一地指向与一个报头段相对应的数据段的缓冲器列表
报头指针的装置；

用于将标志存于内部缓冲器列表的装置；

10 用于将缓冲器列表报头指针存于内部缓冲器列表的装置。

22.根据权利要求 18 的用于组合的系统，其特征在于通信栈是在一
个虚拟电信寻址方式（VTAM）系统中实现的通信栈。

15 23.一个用于将大量数据段和大量报头段组合成数据流用以在网络
通道中传送计算机程序产品，每个数据段唯一地对应于一个上述报头
段，上述网络通道一般要求每个上述数据段和与其唯一对应的上述报头
段相邻，使得数据段及与其对应的报头段有效地在网络通道中传输，该
计算机程序产品包括：

计算机可读的存储媒质，该存储媒质具有包括在该媒质的计算机可
读编码装置，该计算机可读编码装置包括：

20 用于将大量报头段拷贝到数据流缓冲器的报头区的计算机指令装
置，该报头段被顺序地存贮到数据流缓冲器的报头区内；

用于将贮存于数据流缓冲器的报头区中的上述报头段和与其对
应的上述数据段联系起来的计算机指令装置；

25 用于产生包含指示在数据流缓冲器报头区中的上述报头段的入口
记录及指示上述数据段的入口记录的数据流缓冲器列表的计算机指令
装置；

其中上述数据段、数据流缓冲器列表及具有报头区的数据流缓冲器
都可以用来集中地形成用于在网络通道中传送的数据流，而不需物理地
移动数据段，也不需要为数据段产生拷贝，直至数据被传送，该数据流
30 具有报头区和数据区，这样，上述报头区中报头段就可以和与其对
应的上述数据区中数据段不相邻地放在数据流中，该数据流形成于在上述网
络通道中传输该数据流的时候。

邻。但是，许多通信系统为了维持最佳性能，在存储器地址方面都有所限制。

作为例子，在存储器寻址方面的限制有对寻址“范围”条件的限制。比如，有 2k byte 范围的寻址条件，在这种情形下，假设有个字长 50byte 的报头存储在地址 1000 处，相应的长度为 20k byte 的数据存储在地址 5000 处。缓冲区列表的入口记录应当分别指到报头和数据。在这个通信系统中可能存在的限制如下：指向起初数据的指针的缓冲区列表必须指向长度为 20k byte 的数据块，并且列表的增量为 2k byte，这样，20k byte 长的数据段便对应于 2k byte 的地址范围。这种范围要求的例外情况只有：第一个入口记录不必有 2k byte 长，而且不必从 2k byte 的边界上开始入口记录，但必须在 2k byte 的边界上结束；最后的入口记录不必有 2k byte 长，也不必在 2k byte 的边界上结束，但必须在 2k byte 的边界上开始。如果报头是和与其对应的数据段分开存储的，则需要为报头段落准备一个单独的缓冲区列表，因为对每个不同的地址区域都要有个单独的缓冲区列表。

在上面的例子中，如果要在通信系统中传送两个段落，其中每个段落都包含一个报头和相应的数据，若要避免移入数据而形成具有 4 个入口记录（如报头 1 号，数据段 1 号，报头 2 号，数据段 2 号）的单个列表，而不违反缓冲区列表地址的规则（即 2k byte 范围的限制），要完成这项工作通常是不可能的。

一种解决这个问题的已有方案是，将所有的报头和数据段拷贝到单个发送缓冲区中，这样，发送缓冲区中包括：报头 1 号；数据段 1 号；报头 2 号；数据段 2 号；……报头 n 号；数据段 n 号。这个单独的发送缓冲区的信息便在信道中传送。

考虑到通信系统寻址方面的限制，为了减少数据移动的另一种方案是，为每个信息段落提供单独的缓冲区列表，信息段落中包含将在信道中传送的报头及数据。这样，就不会有数据的拷贝或移动的情况出现了。但是，这会造成用于在信道中起传输作用的存储器及缓冲区空间的过多的浪费，还可能发生由于没有数据块的出现而过多地占有输入/输出进程。

最近，还出现了一种方案，它试图不用移动或拷贝数据、不用准备用于信道传输的数据来完成数据的分块，方法是，不管报头的长度，将

其拷贝到和相应的数据段分离的 2k byte 的存储区里。这样，包含报头的存储区也会有大量的“垃圾”数据。报头和这些“垃圾”数据一起在信道中被传输，并且是和数据段分开传输的，这样就可以保持地址范围的一致性。

5 然而，不幸的是，这些已有的方案在遵从通信系统存贮方面的限制的前提下，要解决数据的拷贝和移动的问题，仍显不足。比如，这些方案要求有附加的缓冲区列表，而这便增加了写操作。而且，为了遵守地址范围的限制，这些方案要产生和传送大量的“垃圾”数据以保持地址范围的一致性，这样，存贮空间就被浪费了。更进一步，传送“垃圾”
10 数据都会对节点和通信系统的性能造成负面影响。

还有，已有的方案可能要求通信栈在运行过程中获取新的存贮单元或缓冲区，无论数据段或报头的长度是多少。这样却不仅会由于数据段长度可能与相对较大的数据块的整个长度无关而造成存贮资源的过多的浪费，而且还会由于有大量的数据拷贝和移动而对通信栈的性能产生
15 极大影响。

综上所述，本发明的一个目的就是提供用于在通信系统里管理相对较大的数据块的发送，同时又增进通信系统的性能的方法、系统和计算机程序。

本发明的另一个目的是通过对数据分块，同时减少数据的移动和拷贝来提供用于管理在通信系统中数据的发送的方法、系统和计算机程序。
20

根据本发明，系统提供这些和另外一些目的来增加与在节点间传输大的数据块有关的通信系统的性能，特别是大量的数据段在信道中传输之前，由系统来增加处理它们的能力。

25 在信道中发送数据段之前，管理大量数据段的处理是这样来完成的，标明已形成成包含控制或最初数据的数据流部分，同时继续完成数据分块以避免存贮器资源的浪费。在通信栈的数据链路控制层中，控制或最初的数据拷贝到数据流的报头区内，指向数据流缓冲区中控制或最初的数据的数据流缓冲区列表入口记录也形成了，同时，指向应用数据
30 或静态数据段的数据流缓冲区列表的入口记录也产生了。应用数据段只由数据流缓冲区列表的入口记录所指示。这样一来，就不会产生应用数据的移动和拷贝。

例如，假如采用一个数据流缓冲区列表能在信道中传送的最大数据（比如报头和应用数据）长度是 64k byte。这 64k byte 的数据流可以分为数据流报头区和数据流数据区，其中报头区占 4k byte，数据区占 60k byte。报头区包括由通信栈产生的多个报头，同样，数据区也包括由通信栈产生的多个应用数据段。每个应用数据段与数据流报头区内某个特定的报头相对应。这样，该项发明允许在发送信息时形成数据流，数据流包含报头区和数据区。报头区和数据区形成单个的线性比特数据流，这里的报头段和数据段在数据流中不是相邻地包含在数据流中，而远处的节点则可以从中构成成网络数据单元。

本项发明源于认识到，在通信系统信道中传送数据时，为了减少数据流的存贮资源的浪费，需要将数据合成一个单独的数据流（比如分块）。而且，又认识到通信系统会在存贮地址和存贮范围方面有限制，而这间数据分块发生了矛盾，不仅如此，还认识到在遵守各种各样的通信系统的限制的前提下，数据分块会造成数据的过多的移动和拷贝。如果这种报头和数据段的移动或拷贝被消除了或至少减少了，那么，系统的性能将会得到增强。

这样，根据本发明，准备在通信系统的信道中发送信息时，数据就在通信栈中被系统的 I/O 接口处理，方法是利用数据分块在数据链路控制层中形成“虚拟”的数据流，同时减少数据的移动和拷贝，并遵从存贮范围的限制。在数据链路控制层接收到从通信栈中更高一层传来的多重数据段及相应的报头段，这些多重报头段被拷贝到报头区，这样，报头段就被顺序地存入了报头区。有指向报头区中报头段的入口记录的数据流缓存区列表也产生了，指向从更高一层接收到的数据段的数据流缓存区列表入口记录也存在于数据流缓存区列表之中。在被发送之前，数据段在数据链路控制层中不是物理地移入或拷贝到数据流的数据区中，而是仅仅在系统 I/O 接口直接地拷贝到通信系统的信道中。结果，在通信栈的发送前，只产生一个“虚拟”数据流，并减少了数据移动与拷贝。

实际通过信道发送的数据流由系统输入/输出接口在发送时通过使用数据流缓冲器列表记录以访问存于数据流报头区的报头段及将这些报头段直接写入信道来产生。在所有由数据流列表指示的报头段被写入信道之后，利用相应的数据流缓存区列表入口记录，将数据段（还未被

拷贝)直接从它们原来的存储位置写到通信系统的信道中,因而就减少完成数据分块时紧接着要做的数据拷贝。结果产生了一个具有“带外的”(Out-of-band)报头(即数据流中与相应的数据段不是相邻的报头)的数据流。

5 数据流的报头区是只用来放控制或最初的数据的。这个控制或最初的数据有时又被称作“非相邻的协议数据单元报头”或者是“非相邻的 PDU。”网络协议报头之类的最初的数据与非相邻的 PDU 报头一起,但不必是相邻的,存在于数据流的报头区中。这可有象任何单个的非相邻的 PDU 报头所要求那么多的到数据流的指向标记,以保持数据的联系。
10 只要数据区的控制状态被保持着,在单个或多个数据流中这种联系也就可以被保持着。

15 数据流的数据区总是开始于由数据流报头区末尾指示的预先设置的页面边界。这就会对数据区的开头造成一个固定的偏移。根据所采用的当前的最大缓存区的长度,数据可以是任意长度。这样,任意一条要传送的信息,可以填充到任意大小的数据流中,直到最大的接收缓存区长度减去报头区长度的大小。

20 最后一点,数据链路控制层也通过系统 I/O 接口接收在通信系统的信道中传送的真实的数据流,并且分解数据流,以便将各个报头与其相应的数据段重新联系起来。一旦从在信道中接收到的数据流中分解出一个报头段及与其相应的数据段,那么,它们便被依次送入通信栈的更高的
25 一层接受处理。这个分解过程一直要持续到从信道中接收到的整个数据流都被分解为单个的报头及其相应的数据段。并且每个报头段及其相应的数据段被依次送入通信栈的更高一层之后。

图 1 是一个具有在直接接触的环境中两个主计算机的“应用到应用”通信网络的框图。

图 2 是描述了由两个主计算机及一个网络控制器组成的“应用到应用”通信网络的组成部分的框图。

30 图 3 是描绘了一个通信栈的网络协议层的框图,它源自于遵循“系统网络结构”的计算机软件;该软件运行于如图 1 所示的一个“应用到应用”的通信网络中的一个特定的节点。

图 4 是一个框图,它描述了用于能完成图 3 所示 SNA 构架的通信系统的 VTAM 系统的主要组成部分。

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.